Dissociable effects of visual crowding on the perception of colour and motion

John A. Greenwood & Michael J. Parsons
Experimental Psychology, University College London

Corresponding author: John A. Greenwood
Email: john.greenwood@ucl.ac.uk
Website: http://eccentricvision.com

Summary

Peripheral vision is limited by crowding[1-3], the deleterious effect of clutter that disrupts the recognition of features including orientation[4], colour[5,6], motion[7], and depth[8,9]. Prior research is equivocal on whether this reflects a singular process that disrupts all features simultaneously or multiple processes that affect each independently[5,10-12]. We examined crowding for colour and motion, two features that allow a strong test of feature independence. 'Cowhide' stimuli were presented 15 degrees in peripheral vision; observers reported either the target hue (blue/purple) or its direction (clockwise/counter-clockwise from upwards). We first established that both dimensions show systematic crowded errors (predominantly biased towards flanker identities) and selectivity for target-flanker similarity (greater crowding with similar elements). The multiplicity of crowding was then tested with observers identifying both features: a singular object-selective mechanism predicts that when crowding is weak in one dimension and strong for the other that crowding should be all-or-none for both. In contrast, when crowding was weak for colour and strong for motion, we found that errors were reduced for colour but remained for motion, and vice versa with weak motion and strong colour crowding. This double dissociation follows the predictions of independent crowding processes, suggesting that crowding disrupts vision in a feature-specific manner.

Results

Crowding disrupts our recognition of objects over large regions of the visual field[1-3]. It does so by causing a systematic change in the appearance of target objects[13-15], particularly in the far periphery[16] where targets are induced to appear more similar to surrounding flanker elements. Crowding effects have been observed for a range of visual features including orientation[4], position[13], spatial frequency[10], size[5], colour[5,6], motion[7], and depth[8,9]. Within these dimensions, crowding is also modulated by the similarity between target and flanker elements – differences in features including orientation, colour, and depth can all reduce the strength of crowding[4,6,12]. Given the distributed way in which these features are likely to be processed throughout the visual system[17,18], can a single process produce this multitude of effects? Most models of crowding implicitly assume that crowding occurs via a single combined process within the visual system, particularly for higher-order approaches where crowding derives from attention[19-21] or grouping[22]. If crowding were instead to operate independently across distinct visual features, then these effects could potentially involve an array of distinct neural substrates with varied mechanisms.

A key prediction for a combined crowding process is that a release from crowding in one feature domain (e.g. colour) should release crowding for other features (e.g. orientation) at the same time. Accordingly, target-flanker differences in colour or contrast polarity can reduce crowding for judgements of spatial form[12], while differences in orientation can improve crowded judgements of position[11]. Others have however found that errors in the judgements of spatial frequency, colour, and orientation follow a combination of independent and combined error patterns[10]. This discrepancy may arise from the specific features used in each study. Here we sought to investigate whether crowding is combined or independent for judgements of colour and motion – arguably the two features with the clearest separation in the visual system[17,18]. We conducted three experiments with these features, each using cowhide-like stimuli[23,24] presented 15 degrees in the upper visual field. In Experiments 1 and 2 we measured the nature of crowding in each feature dimension separately – in particular, the target-flanker combinations that produce strong vs. weak crowding and the nature of the errors that result (i.e. whether the errors are consistently systematic in nature, as observed for orientation[15] and position[13] previously). This allowed us to identify a set of conditions where crowding had either a strong or weak effect on judgements of each feature independently, and to predict the appearance of target stimuli under these circumstances. In Experiment 3 we then examined the effect of these manipulations on conjoint judgements of both features in order to test the independence of crowding across these dimensions.

In Experiment 1, observers viewed moving cowhide stimuli and reported the direction of movement (clockwise/counter-clockwise of upwards) for either an isolated target (Figure 1A, left panel) or a target with flankers positioned above and below at a centre-to-centre separation of 2.25° (Figure 1A, middle and right panels). Flankers moved in one of sixteen directions ±0-180° from upwards. Data from an example observer are displayed in Figure 1B, where the proportion of counter-clockwise (CCW) responses is plotted as a function of the target direction (depicted schematically on the x-axis), along with the best-fitting psychometric function for each condition.

For the example observer in Figure 1B, uncrowded judgments (blue points) transitioned rapidly from predominantly clockwise to counterclockwise for directions around upwards (0°). The best-fitting psychometric function is accordingly steep and crosses 50% CCW near to upwards, indicating high sensitivity and low bias. In the presence of flankers moving upwards (+0°; yellow points), performance declined and the slope of the psychometric function is accordingly shallower, with larger directional offsets required to reliably report the direction of the target. Performance remained unbiased however, with a PSE that crosses 50% at the point of upwards motion. In contrast, flankers moving 30° CCW of upwards (red) induced a strong bias towards CCW responses, resulting in a leftwards shift of the psychometric function in addition to the shallower slope. The opposite bias can be seen for CW flankers (purple). These assimilative errors can thus be captured using the PSE, while the impairment in performance across all three of these crowded conditions is shown via threshold values.

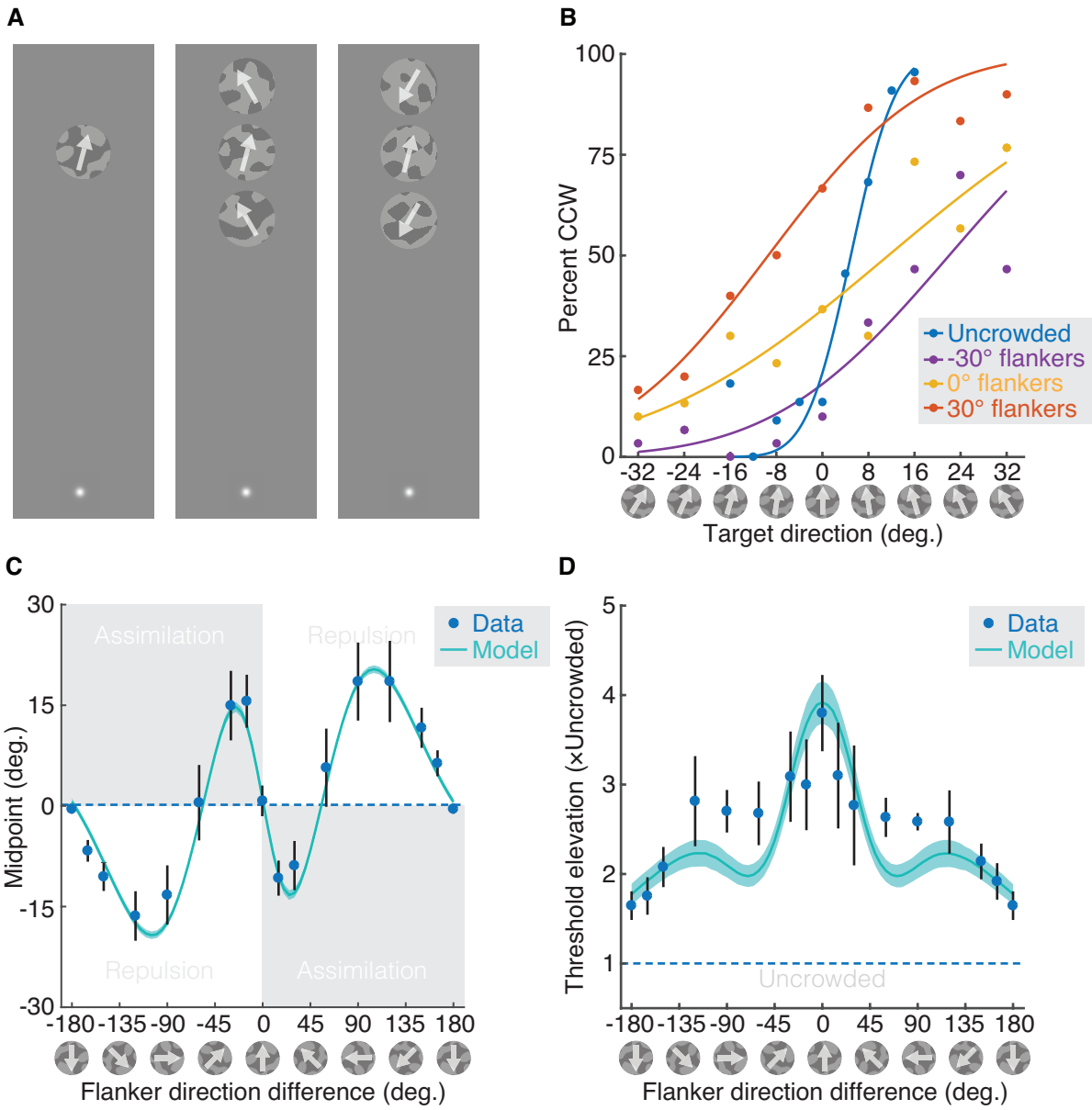


Figure 1. The effect of crowding on motion perception. **A.** Left panel: An example uncrowded stimulus. Observers judged the direction (CW/CCW of upwards) of an isolated 'cowhide' stimulus. Middle: An example crowded array where the target (middle, moving CW of upwards) is presented between two flankers moving 30deg CCW of upwards. Right: An example crowded array with flankers moving 150deg CCW of upwards. **B.** Example data and psychometric functions for observer YL. Data are presented as the proportion of trials with a CCW response, plotted as a function of target orientation. Uncrowded data (blue) show a sharp transition from predominantly CW to predominantly CCW responses with a PSE near upwards (90°). Crowded conditions all show shallower slopes, indicating worse performance, with PSEs indicating no bias for upwards moving flankers (yellow), and assimilative biases with a decrease in CCW responses for the 60° flankers (purple) and vice versa for the 120° flankers (red). **C.** Midpoint (PSE) values averaged over all 6 observers (blue points with error bars showing SEM), plotted as a function of the flanker direction. The mean output of a population model of crowding is shown (green line) with the 95% range of values from 1000 simulations depicted around it. **D.** Threshold elevation values (crowded divided by uncrowded performance), for the same conditions, plotted as in panel E.

Psychometric functions were fit separately for each flanker condition and observer. Mean PSE values across observers are plotted as a function of the target-flanker difference in Figure 1C. On average, upwards moving flankers (0°) did not induce any bias, as with the example observer above. Flankers moving slightly clockwise (e.g. -15°) induced a positive shift in PSE, indicating an increase in clockwise

responses. These assimilative errors were mirrored for small target-flanker differences in the counter-clockwise direction. Larger target-flanker differences (e.g. ±90°) instead induced a repulsive shift in the PSE, indicating that the perceived target direction was biased away from that of the flankers. Further increases gave a reduction in bias, with downwards flankers inducing no bias on average.

To quantify performance, we divided crowded thresholds by the uncrowded baseline to obtain the threshold elevation scores presented in Figure 1D, where a value of 1 indicates performance as precise as that for an uncrowded target (dashed line). The greatest threshold elevation occurred with upwards-moving flankers, with a steady decline in threshold elevation with flanker directions away from this. Threshold elevation was at its lowest point with downwards moving flankers, though the elevation remained clearly above 1 for all observers.

To better understand the mechanisms underlying these patterns, and to work towards predictions for combined vs. independent models of crowding, we developed a population-coding model to simulate these crowding effects. Given the systematic nature of these errors, the most plausible existing models are those based on averaging[13,15] or substitution[25] processes. A more general approach has recently been shown to account for both averaging and substitution errors by modelling crowding as the combination of population responses to the target and flanker elements[26]. Although these approaches are focussed on the assimilative errors in crowding, repulsive errors have previously been simulated by weighted averaging models via the use of negative weights[27]. Here we sought to incorporate these errors into a population-based approach using inhibitory interactions between the target and flanker population responses, similar to models of the tilt illusion[28] and to surround effects in cortical area MT/V5[29]. We therefore simulated a population of direction-selective neurons with a Gaussian profile of responses to direction, similar to those reported in V1[30] and MT/V5[31] (see supplementary materials for parameters and further details). The population-response profiles to both target and flanker directions were combined according to a weighting field[26] (see Supplementary Figure 1). Where previous studies[26] have modelled the weighting field as a function of target-flanker separation, here we model its variation as a function of target-flanker similarity (as suggested by prior work[16]). The weighting field had both excitatory and inhibitory components to allow the production of assimilative and repulsive effects, respectively. Both components declined with the increase in direction difference between target and flanker elements (rather than with target-flanker distance as in prior models[26]) in order to simulate the effects of target-flanker similarity[4,6,12,16].

Best-fitting parameters were obtained with a least-squares fit of the model to the mean data, with the resulting simulated bias values shown in Figure 1C (green line). The model follows the increase in assimilative bias with small target-flanker direction differences because the overlap between target and flanker population responses induces a summation of the signals. The model also simulates the rise and fall of repulsion with larger differences when the target is inhibited by the negative component of the flanker response (see also Supplementary Figure 1). Similarly, threshold elevation values (Figure 1D) show the greatest elevation for small direction differences, and a gradual decline on either side.

In sum, for the crowding of motion, the greatest threshold elevation was found for small target-flanker differences, where the errors induced were assimilative in nature. Larger target-flanker differences produced a reduction in threshold elevation and errors of repulsion from the flanker directions, with further reductions in bias and threshold beyond this. A population-based coding model with Gaussian tuning for direction and inhibitory surrounds captures this pattern well.

We next examined the effect of crowding on judgements of hue in Experiment 2. Here observers identified whether the target was blue/turquoise or purple/pink (Figure 2A, left panel). When present, flankers had one of twelve different hue angles in DKL colour space[32-34], which differed from the reference hue by values from 0° to ±180° (see Figure 2A, middle and right panels). Data from an example observer are displayed in Figure 2B, where the proportion of pink responses are plotted as a function of the target hue (again depicted on the x-axis). Here, flankers with the same hue as the reference boundary (0°; yellow points) did not induce a bias, though the slope is clearly increased

relative to the uncrowded condition (blue points). Flankers with a hue angle difference of +15° (a shift towards pink hues; red points) produced both an increase in the slope of the psychometric function and a shift in the function, indicative of assimilative errors. The opposite is seen for -15° flankers, where a shift towards blue hues in the flankers also induced assimilative errors (purple points).

Mean values of the PSE across all flanker hue conditions are plotted in Figure 2C. As with motion, flanker hues at the decision boundary (0°) induced no bias on average. Flankers with clockwise/negative hue angles, blue or green in appearance, again induced a positive shift in PSE, indicative of an increase in 'blue' responses. Larger target-flanker differences (approaching brown/yellow in appearance) gave a reduction in this bias. These assimilative errors were mirrored for flankers with counter-clockwise/positive hue angles, ranging from purple and pink through to red in appearance. Unlike motion however, no repulsive shifts in PSE were observed – larger target-flanker differences simply gave a reduction in assimilative bias.

Mean values of threshold elevation are presented in Figure 2D. Although overall threshold elevation values are lower than for direction, the pattern of data is broadly similar, with the greatest threshold elevation for small target-flanker differences and a decrease in crowding strength with increasing difference. At the end of the range, flankers with the opposite hue to the reference boundary in DKL space (yellow/brown) did not produce threshold elevation relative to uncrowded performance.

As with direction, we modelled this pattern of crowding using a population-based approach where each detector was selective for a distinct hue angle in DKL colour space[34,35] with a Gaussian profile of selectivity (suggested by psychophysical results[36] and used here for ease of comparison across dimensions). Because repulsion was not observed for hue under these conditions, the inhibitory surrounds for detectors and the weighting field were set to zero. Bias values obtained from the model with the best-fitting parameters are plotted in Figure 2C (green line). The model again captures the high degree of assimilative errors with small target-flanker differences in hue, and the decline of these errors at larger differences. Threshold elevation values are similarly well captured (Figure 2D), with a high degree of impairment with small target-flanker differences and a reduction in this effect at larger differences.

To summarise the effects on colour crowding, small target-flanker differences in hue induce strong performance decrements and assimilative errors, with similarly assimilative errors at larger differences despite the reduction in crowding strength. The overall pattern is broadly similar to that seen for motion, though with an absence of repulsive errors and an overall reduction in threshold elevation. A population-based model again captures this pattern well.

Despite some differences then (notably in the presence or absence of repulsive errors), in both of these dimensions it is clear that crowding produces systematic patterns of errors that are predominantly assimilative for small target-flanker differences. Both dimensions also show selectivity for target-flanker similarity such that crowding is strong when the target and flanker were similar and weak when they differ. Armed with this knowledge, we can now make predictions for paired judgements of both colour and motion. Namely, when crowding is strong for one dimension (with small target-flanker differences, e.g. in direction) and weak for the other (with larger differences, e.g. in hue), independent crowding processes allow for assimilative errors to occur for the dimension with strong crowding and for no errors to occur in the other. In contrast, a combined/singular mechanism predicts that crowding must be all-or-none: if crowding is weak in one dimension and strong in the other, then it must either be released for both (i.e. no errors should occur) or the errors must persist in both dimensions. Experiment 3 was designed to test between these alternatives.

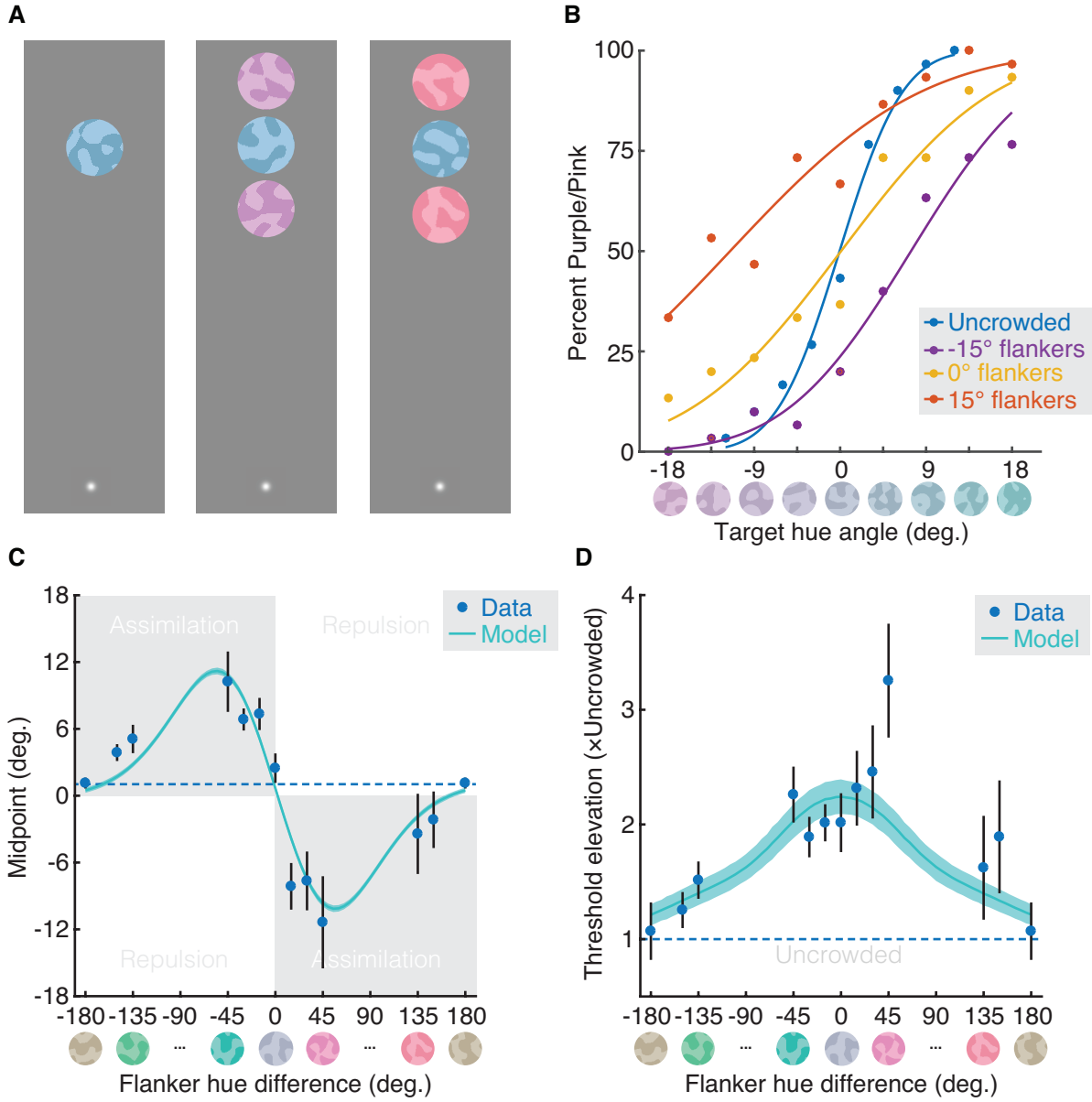


Figure 2. The effect of crowding on colour perception. **A.** Left: An example uncrowded stimulus. Observers judged the hue (blue/purple) of an isolated 'cowhide' stimulus. Middle: An example crowded array where the target (middle, +16° blue in DKL space) is presented between two flankers 15° from the blue/purple boundary in DKL space (purple). Right: Example crowded array with flankers 135° from the blue/purple boundary (pink/red). **B.** Example data and psychometric functions for observer YL. Data are presented as the proportion of trials with a 'purple/pink' response, plotted as a function of target hue. Uncrowded data (blue) show a sharp transition from predominantly blue to predominantly pink/purple responses with a PSE near the boundary (0°). Crowded conditions all show shallower slopes, indicating worse performance, with PSEs indicating no bias for flankers coloured with the boundary value (yellow points), and assimilative biases with a decrease in pink/purple responses for the 15° flankers (red) and vice versa for the -15° flankers (purple). **C.** Midpoint (PSE) values averaged over all 6 observers (blue points with error bars showing SEM), plotted as a function of the flanker hue. The mean output of a population model of crowding is shown (green line) with the 95% range of values from 1000 simulations depicted around it. **D.** Threshold elevation values (crowded divided by uncrowded performance), for the same conditions, plotted as in panel E.

In Experiment 3, observers made conjoint judgements of the direction (CW/CCW of upwards) and hue (blue/pink) of the target cowhide stimulus. We conducted these judgements for targets in isolation and for three crowded conditions. In the first, crowding was strong in both feature dimensions, with small target-flanker differences in both direction and hue. In the second, crowding was strong in direction and weak in hue, with a large colour difference between target and flanker elements. The

third was the converse, with weak crowding for direction (large direction differences) and strong crowding for hue. In each case, targets were presented with one of two directions and one of two hues, and likewise in the flankers (see Methods for values). As a result, there were four possible combinations of target and flanker elements, grouped according to their agreement relative to the 2AFC decision boundary for each dimension: target and flanker elements either matched in sign relative to the decision boundary for both dimensions (e.g. clockwise moving target and flanker elements, with a blue hue in each), differed in colour (e.g. a pink target with blue flankers, all moving clockwise), differed in motion, or differed in both. Note that these distinctions ignore the precise values of direction and hue (e.g. a target moving 8° clockwise of upwards has the same sign as a flanker moving 15° clockwise).

With these four combinations of target and flanker elements and three crowding-strength conditions, we can make specific predictions for combined vs. independent mechanisms of crowding. For a combined mechanism, crowding is a singular process that operates in an all-or-none fashion. Thus, particularly in the 'both differ' condition, the combined mechanism predicts that errors should be made either in both dimensions or in neither. In the case where crowding is strong in both dimensions, this is likely to hold true. However, when motion is weak due to large target-flanker differences, the combined mechanism predicts either that motion errors should persist (due to strong crowding for colour) or that the colour errors should be released as well. In contrast, the independent mechanism predicts a reduction in motion errors and the persistence of the colour errors, because crowding can be reduced in one dimension without affecting the other. The same is true with the opposite condition, where the combined mechanism predicts an all-or-none pattern while the independent mechanism predicts motion errors in the absence of colour errors.

In order to quantify these predictions, we developed two versions of the population-coding model, with the majority of parameters fixed by the results of Experiments 1 and 2. The independent model consisted of population responses to motion and colour, generated for both target and flanker elements and combined according to separate weighting fields for both motion and colour. These separate weighting fields allow for crowding to occur for one dimension (with small target-flanker differences, e.g. in colour) and not in the other (when target-flanker differences for a given dimension are larger, e.g. in motion). For simplicity and ease of comparison to the combined model, only the positive weighting field and excitatory population responses were used in these models. The combined model was otherwise identical except for the use of a common weighting field for the two dimensions. Thus, if crowding occurs for one dimension it necessarily occurs for both, and vice versa. Further details are provided in the methods below, with the best-fitting parameters outlined in the supplementary materials.

When the target was uncrowded, observers were successfully able to identify the target direction in 87.71 ± 3.29% (mean ± SEM) of trials, and the target hue in 93.96 ± 1.76% of trials. Results from the three crowded conditions are presented in Figure 3. Panel A shows the mean responses when crowded with target-flanker differences that were small in each dimension (which, as in Experiments 1 and 2, gives strong crowding for motion and colour). Here, when both dimensions matched (red point), percent correct values were high for both dimensions. This is predicted by the assimilative nature of the flankers in each case – even if crowding occurred, it would still pull responses towards the correct direction/hue. In the *motion differs* condition, the flankers had the opposite sign to the target in direction but not colour, and a predominance of motion errors resulted. This again is predicted by assimilative errors for direction, with either no effect on colour or assimilative errors which pull the target hue in the same direction as the flankers. The opposite occurred for the *colour differs* condition, with the small colour difference inducing assimilative colour errors. Finally, in the *both differ* condition, the combined effects of assimilation in each domain induced a preponderance of errors in both dimensions.

Figure 3B shows results from the weak motion (strong colour) crowding condition. Note here that the agreement in target-flanker sign is determined relative to the upwards decision boundary, meaning that flankers moving -165° from upwards (clockwise-downwards) have the same sign as a target

moving -15° (clockwise). Here, in the *both match* condition, responses were predominantly correct on both dimensions, as before. In this case however, in the *motion differs* condition, there is now a reduction in crowding, and a corresponding increase in correct responses for motion in this condition, while the matched target-flanker colours leave the colour responses at an equivalently high level – responses are now predominantly correct in both dimensions. As before, in the *colour differs* condition the small target-flanker difference induces assimilative errors, while the matched sign of the target and flankers in direction could reflect either assimilative errors or correct target recognition. Crucial is the *both differ* condition: here, the large difference in target-flanker direction coupled with a small difference in target-flanker colour produced a predominance of colour errors. In other words, crowding is strong for colour whilst also being weak for motion in the same stimulus.

The reverse pattern can be seen in Figure 3C, which plots results from the weak colour (strong motion) condition. The *both match* condition again produced a high proportion of correct responses. In the *motion differs* condition, the small target-flanker differences in motion induced assimilative errors (as in panel A when crowding was strong in both dimensions), with a high rate of motion errors and low rate of colour errors (which again could be due to either assimilative errors or correct target recognition). Here in the *colour differs* condition the large colour difference induced a reduction in crowding for hue judgements, with a corresponding increase in correct responses on colour and the matched target-flanker signs for motion leading to correct responses in both dimensions. Finally, the *both differ* condition again revealed a dissociation – the large differences in target-flanker colours coupled with a small difference in their directions produced a predominance of motion errors. Here too, crowded errors can occur in one dimension and not the other.

Simulations from the independent and combined crowding models are shown in Figure 3D for the condition where crowding in both dimensions was strong. Given that the population model can reproduce the results of Experiments 1 and 2, both models are similarly able to capture this pattern of results. The independent model does well because the probability of crowding (via the two weighting fields) is high for both dimensions. The combined model similarly does well because crowding is all-or-none and the target-flanker differences are small in both dimensions, leading to a high degree of crowding in each case. Both models can therefore produce crowding effects when the strength in each domain is maximised.

Figure 3E shows the best-fitting model outputs for the weak motion (strong colour) condition. Here the independent model successfully captures the pattern of performance because the separate weighting fields for the two dimensions means that crowding can be decreased in one domain and not the other, leading to a preponderance of colour errors in the *both differ* condition. However, the best-fitting parameters of the combined model also lead to a similar pattern of performance here. This can be attributed to two factors – the broad weighting field for colour crowding with a low peak strength (required to capture the results of Experiment 2, unlike the more narrowly tuned and higher-peaked motion weighting field from Experiment 1; as shown in Supplementary Figure 1) and the broad tuning of detectors for colour in the population. As a result, when motion crowding was reduced in strength the weighting of flankers became driven by the target-flanker differences in colour, giving some improvement in performance (relative to when crowding was strong in both domains) due to the lower weights for colour, though the broad tuning of colour detectors still produced a predominance of errors in this domain due to the overlap in responses to the target and flanker signals in the population (whereas the more widely separated motion distributions allowed for a clearer separation). In other words, even a combined model where crowding is either on for both dimensions or off for both can at times appear to follow an independent pattern of errors.

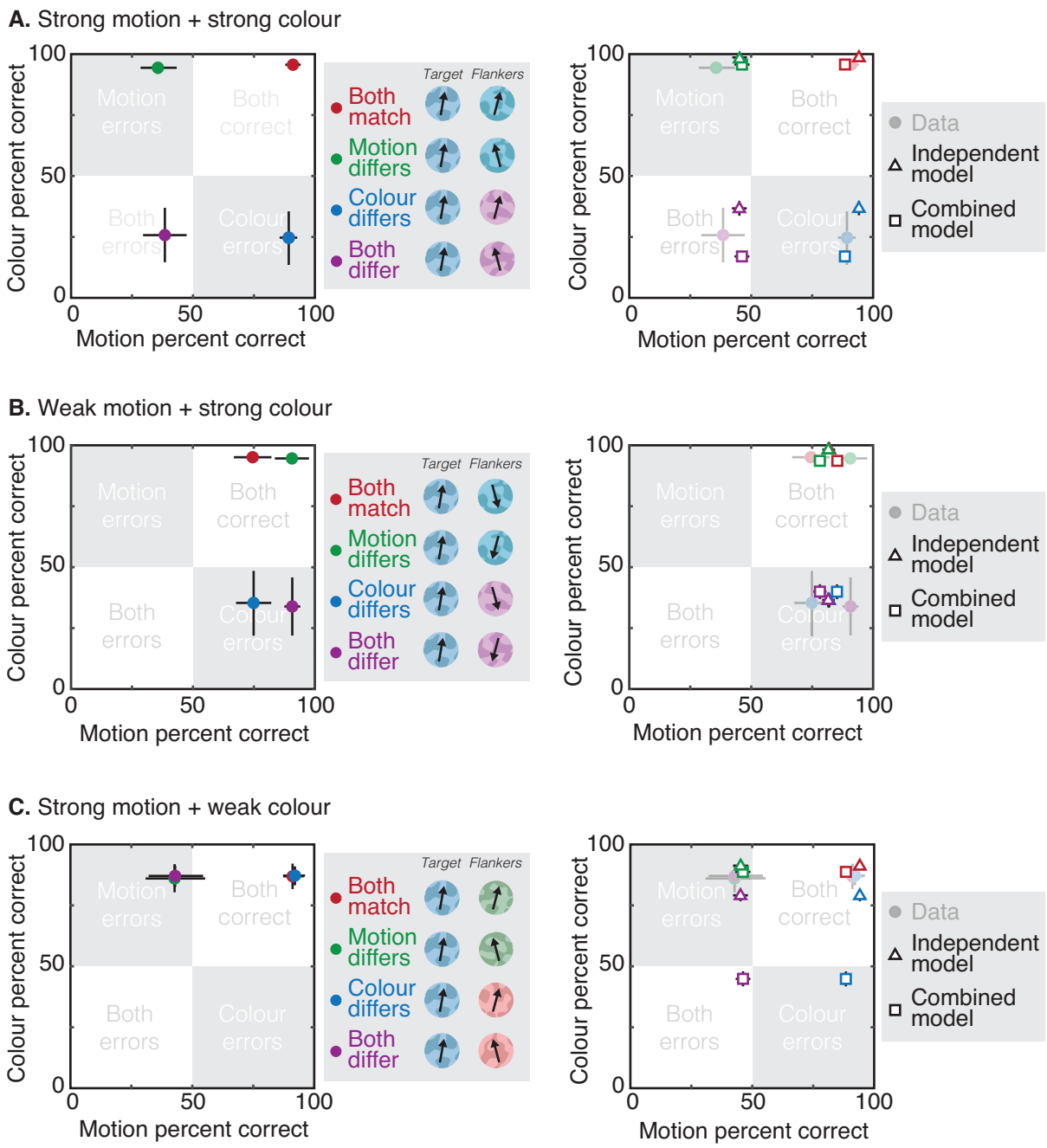


Figure 3. Results from the dual crowding experiment. Each panel plots the mean proportion of responses (n=6) that were correct for the target direction of motion on the x-axis and the target hue on the y-axis. With this we can categorise responses as being correct on both dimensions, wrong on motion or colour, or wrong on both, as indicated by the quadrant demarcation. We further divide conditions into 4 target-flanker match cases: where the 2AFC sign in each dimension matches for both, where the motion differs, the colour alone differs, or both differ. Data points are shown as the mean ±1 SEM. **A.** Left panel: Responses when flankers were selected to produce strong crowding on both dimensions (i.e. with small target-flanker differences in each). Right panel: Model simulations for the independent and combined population models in the both strong condition. **B.** Left: Responses with a large target-flanker difference for motion (which induces weak crowding) and a small target-flanker difference for colour (which induces strong crowding). Right: Model simulations for the weak motion (strong colour) condition. **C.** Left: Responses with a large target-flanker difference for colour (weak crowding) and a small target-flanker difference for direction (strong crowding). Right: Simulations for the weak colour (strong motion) condition.

Model outputs diverge however in the weak colour (strong motion) condition, shown in Figure 3F. Here the smaller target-flanker differences for direction lead the combined model to be driven by the

high-peaked weighting field for motion. In this case crowding occurs strongly for both dimensions, leading to the erroneous prediction of colour errors when *colour differs* and errors in both dimensions when *both differ*. Thus, although the best-fitting parameters are sufficiently narrowly tuned to allow reproduction of the errors when motion crowding was weak, the same parameters necessarily fail in this condition. In contrast, the independent model is able to successfully reproduce the pattern of data in this condition because crowding can be reduced for colour and remain strong for motion, leading to a preponderance of motion errors in the *both differ* condition. Our results therefore offer clear support for an independent mechanism for crowding in these two domains.

We have further explored a range of variations in these models, with a focus on whether alternative versions of the combined model could improve its ability to capture performance. In the supplementary materials, we report that the least-squares error for model fits can be slightly reduced in the independent model with the restoration of the negative weighting field and flanker components, though the same manipulation increases error for the combined model (Supplementary Figure 2). Further variations of the combined model similarly do little to improve its performance – selecting the minimum probability for crowing in both dimensions (instead of the maximum in the above modelling), using a common weighting field for both dimensions, and using shared population characteristics all produce worse fits than the independent model (Supplementary Figure 3). In other words, the precise mechanisms of these models do not matter greatly – the effect of crowding on colour and motion is better explained by independent processes for the two dimensions than by a single combined mechanism.

Discussion

Our results demonstrate that independent crowding effects disrupt the perception of motion and colour. In our first two experiments, crowding with small target-flanker differences impaired both direction and hue judgements with strong assimilative effects and high threshold elevation, while larger target-flanker differences gave a reduction in threshold elevation. This allowed us to select specific combinations of target-flanker values where crowding could be strong or weak in one or both dimensions (by decreasing or increasing target-flanker differences, respectively) whilst observers made conjoint judgements of the target direction and hue. In this third experiment, when crowding was weak for motion, crowding for colour remained strong. Similarly, a reduction in the crowding of colour did not affect errors for the target direction. A population-coding model of crowding could account for these results with weights applied to the flanker values in each feature dimension that were independently determined by separate 'weighting fields'. Models where crowding operates as a singular all-or-none process (with the same strength of crowding used for both dimensions at the same time) failed to replicate the observed pattern.

Further dissociation between colour and motion can be seen in the differences in the pattern of midpoint values with large target-flanker differences in Experiments 1 and 2. Although the crowding of motion led to a repulsion in the perceived direction of moving targets, large target-flanker differences in colour simply reduced the rate of assimilative errors. Our population models could account for these patterns through inhibitory interactions in the motion population, which were not required to produce the pattern of colour biases. This is not to say, of course, that contextual modulations for colour can never be repulsive. In the fovea, repulsive shifts have been reported for the perceived hue of targets surrounded by adjacent elements[37], similar to the repulsion observed for perceived direction in both foveal[38,39] and peripheral[39] vision, though contextual effects for hue in the periphery often tend towards assimilation (e.g. with neon colour spreading[40]). In the domain of orientation, repulsive effects of tilt contrast predominate near the fovea, giving way to assimilation as stimuli are moved into the periphery[16]. It is possible then that the transition from assimilation to repulsion occurs at lower eccentricities for colour than for motion, which would produce a divergence in the perceptual outcomes of crowding for some parts of the peripheral field. Further differences are apparent in our threshold elevation values, which were substantially lower for the crowding of colour than for motion. This is consistent with prior findings comparing colour with domains like orientation and size[5]. We nonetheless reproduced these differences in our model with variations in the strength of inhibitory interactions, and with a lower peak in the weighting field to combine the colour of target

and flanker elements than in the weighting field for motion. In other words, population responses to colour were not only affected by crowding in a different manner to motion (with assimilation alone instead of a mixture of repulsion and assimilation) but with a lower overall magnitude.

This dissociation in the effect of crowding for colour and motion differs somewhat from prior studies using other visual features. It is possible that the specific features chosen matter greatly in this regard. The mixed pattern of independent and combined errors seen with spatial frequency, colour, and orientation[10] could have arisen because colour dissociates from orientation and spatial frequency (as others have recently suggested for crowding[41]), while orientation and spatial frequency are more closely linked. Similarly, the observation of combined errors for orientation and position judgements found using a similar paradigm[11] could reflect the closer relationship between these feature domains[42] than for colour and motion[17,18] (though of course the separation between colour and motion is not absolute[43,44]). More generally, a dissociation between colour and motion has also been observed for visual masking effects[45]. In visual working memory, colour and orientation show a dissociation via independent rates of decay, in contrast with linked spatial dimensions for object height and width[46]. Similarly, although visual search for objects defined by conjunctions of spatial features is typically slow and serial[47,48], conjunctions of motion and form can be found extremely quickly[49], suggesting a separation for some dimensions and not others. A mixture of linked and dissociable effects for different feature combinations may thus be common for processes throughout the visual system.

We have modelled the effects of crowding with a population-coding process, similar to prior approaches[26]. In fact, the dissociable nature of crowding lends itself to this approach – separate populations of detectors selective for motion and for colour could each produce these effects separately, unlike a combined process which would require detectors with a combined selectivity for these features. Nonetheless, the precise mechanism does not matter greatly in terms of the dissociation in performance that we observe. For instance, it is likely that 'texturisation' models[50-52] could perform similarly to our population models in producing the systematic effects observed here for crowding, though repulsion effects are yet to be examined with these approaches, and distinct implementations would at least be required for the temporal and spatial properties of these textures in order to produce the dissociations observed here for moving stimuli.

Our results are however inconsistent with higher-level theories of crowding. For instance, Gestalt approaches[22] argue that crowding occurs whenever the target forms a good Gestalt with the flankers. With these manipulations, crowding is increased when the target becomes part of a pattern including the flankers[53,54], or reduced when the flankers form a shape[55] or a row of elements[56] that excludes the target. In the current context, grouping theories predict that crowding should be strong when the target and flankers are grouped together and weak when they are ungrouped, a process that should apply to the collection of features within the target as a whole. In other words, grouping is necessarily an all-or-none process that is inconsistent with the dissociations we observe here. Our findings are similarly unlikely to be accounted for by attentional theories of crowding[19-21] since the higher-level nature of attentional selection would predict that crowding operates at the level of objects and/or locations rather than specific features. Of course, it is still possible for attention and grouping effects to modulate the strength of crowding – our findings simply demonstrate that these processes are not the key factors in the mechanism of crowding.

The dissociation in these impairments to colour and motion have further implications for the neural locus of crowding. Prior work suggests that crowding may involve multiple cortical sites, with neural correlates reported in V1[57,58], V2 and V3[59], while others have noted similarities between crowding and the properties of V4 neurons[60]. Judgements of crowded orientation reveal effects throughout the cortical hierarchy, with BOLD responses in higher levels altered to an increasing extent[61]. Our current findings suggest that the widespread nature of these effects may reflect the precise features used, i.e. that crowding for dissociable features like colour and motion may rely on distinct neural substrates. Although the search for a neural basis for crowding has tended to focus on the ventral stream[51], similar (but independent) processes may occur within the dorsal stream to give rise to the crowding of motion. Alternatively, this distinction could arise within a single cortical region if the effect of

crowding on colour and motion were to rely on distinct neural subpopulations (as argued in the context of feature binding[62], for instance).

These findings bear some similarity to arguments that crowding operates in a multi-level hierarchical fashion[1,63]. However, these arguments are based primarily on an apparent difference in the selectivity of crowding for faces compared with other stimuli[64], an effect that disappears once the task difficulty with upright and inverted faces is equated[65]. Perhaps similarly, we do observe some differences in the operation of crowding for colour and motion, with repulsion for motion for larger target-flanker differences, whereas colour crowding only produced assimilation. In this sense, there may be subtle differences between the operation of crowding in distinct domains, though the broad selectivity in terms of the strength of crowding was highly similar – for both motion and colour, crowding was strong with small target-flanker differences and weak with larger differences. In other words, wherever crowding occurs, it appears to follow broadly similar principles in each case.

Finally, given the dissociable effect of crowding for distinct feature domains, we may think of crowding as a general property of the visual system, similar to processes like adaptation that occur in a range of domains including orientation, motion and colour through to high-level properties like faces[66-68]. Just as adaptation likely serves to increase processing efficiency and to highlight novel stimuli over time, crowding may perform a similar function by simplifying visual input[14,15] when target-flanker elements are similar and allowing stimulus pop-out when they are dissimilar[69,70]. The selectivity required to produce these changes in appearance is likely to require distinct neural populations (e.g. the changes required to alter the appearance of colour vs. motion), but the general processes are clearly similar in each case (e.g. that effects tend towards assimilation). In sum, we propose that crowding is a generic visual process that operates by pooling the identity of target and flanker objects in a feature-specific manner in order to simplify the appearance of our peripheral vision.

Experimental Procedures

Observers

Six observers (3 male, including the two authors) participated in all 3 experiments. All had normal or corrected-to-normal visual acuity, and scored 17/17 on the Ishihara test[71] indicating normal colour vision.

Apparatus

Experiments were programmed in MATLAB (Mathworks, Inc.) on an Apple Mac Pro using the PsychToolbox[72,73]. Stimuli were presented on a 21" Mitsubishi Diamond Plus CRT monitor with a resolution of 1400×1050 pixels and a 75Hz refresh rate. The monitor was calibrated using a Minolta photometer and linearised in software to give a mean and maximum luminance of 50 and 100 cd/m^2, respectively, and a white point close to the standard CIE Standard Illuminant D65. Maximum luminance values for red, green, and blue were 28.3, 69.5, and 8.1 cd/m^2, respectively. Observers viewed stimuli binocularly from a distance of 50cm, with head movements minimised using a head and chin rest. Responses were given with the use of a keypad, with auditory feedback provided only during practice sessions.

Stimuli and Procedures

In all experiments, target and flanker stimuli were 'cow-hide' elements[23,24], created by band-pass filtering white noise with a spatial frequency cut-off of 1.5 cyc/deg, and rounding the luminance to give two values (light and dark). Each element was presented within a circular aperture with 2° diameter. The advantage of these elements was that the visible contours enabled the percept of motion whilst also allowing the alteration of the surface hues.

The target element was presented at 15° eccentricity in the upper visual field, either in isolation or with one flanker above and one below when crowded. The centre-to-centre separation of target and elements was 2.25° - a value of 0.15× the eccentricity, ensuring strong crowding effects[2,3,74]. Observers fixated on a two-dimensional Gaussian blob with a standard deviation of 4' of arc, presented just below the centre of the screen, and were required to maintain fixation throughout the experiment.

Stimuli were presented for 500 ms, followed by a mask for 250 ms. The mask was a patch of 1/f noise presented in a circular window with a diameter of 4.8° in the uncrowded conditions and 8.5° when crowded, plus a cosine edge that spanned an additional 1° on all sides. The mask was followed by a blank screen with only the fixation point, at which time observers were able to respond.

In Experiment 1, the cow-hide stimuli were presented as grey-scale moving elements where light and dark regions had a Weber contrast of ±0.75. Movement was obtained by generating patches as a long strip of texture that moved underneath the aperture with a displacement of 5.8' of arc per frame every second monitor frame (to allow greater resolution of directional displacements with larger, less frequent step sizes). This gave an effective refresh rate of the stimulus at 37.5 Hz and a speed of 3.6 deg/sec.

When uncrowded, the target element moved in one of 9 equally spaced directions between ±16° around upwards. The greater difficulty in crowded conditions required an increase in this range to ±32°. Observers indicated whether the target moved counter-clockwise or clockwise of upwards using the '4' and '6' keys, respectively, on a numeric keypad. When crowded, flankers moved together in one of sixteen directions relative to upwards: 0°, ±15°, ±30°, ±60°, ±90°, ±120°, ±150°, ±165°, or 180°. With 9 target directions and 10 trials per direction there was a total of 90 trials for each block of the uncrowded condition. Crowded conditions were blocked with two opposite flanker directions (e.g. +15° and -15°) to ensure a balanced likelihood of clockwise and counterclockwise responses. The 0° and 180° conditions were blocked together for consistency. This gave 180 trials for each block of the crowded condition. Each block was performed 3 times, giving a total of 4590 trials per observer, plus practice, completed in 3-4 sessions of 1 hour each including breaks.

In Experiment 2, cowhide elements were static and presented with a range of varying hues. Colours were determined using the DKL colour space[32-34] with a luminance contrast of ±0.3 for the light and dark regions of the cowhide and a colour contrast (saturation) of 0.2 for all colours. Variations were applied solely to the hue angle of each patch. The reference hue angle was determined individually by presenting the test range of hues (from blue/turquoise to pink/purple) and asking observers to indicate the midpoint where stimuli appeared closest to neutral. This gave a reference hue angle of 262.5° for four observers, 262.0° for JG, and 264.0° for CS. In the uncrowded conditions, target hues ranged in 3° increments from the base hue angle, between +12° and -12°. For the crowded conditions, the target varied colour in 4.5° increments between +18° and -18° of the base hue angle. Observers judged whether the target patch appeared blue/turquoise (clockwise in DKL space) or purple/pink (counter-clockwise) using the '2' and '8' keys, respectively. Observers were shown the range of target hues prior to each block of trials, as well as the flanker hues in crowded conditions. When present, flankers had one of twelve different hue angles relative to the base hue angle: ±15°, ±30°, ±45°, ±135°, ±150°, 0° and 180°, tested in blocks that contained opposing directions as above. This again gave 90 trials per uncrowded block and 180 when crowded, giving a total of 3510 trials per observer, plus practice, completed in 3 sessions of 1 hour each, including breaks.

Finally, in Experiment 3, cowhide elements were presented with both motion and colour values. Using values derived for each observer in the first two experiments, we first selected target direction and hue values that gave near-ceiling levels of performance when uncrowded but that were still affected by crowding in the strongest crowding conditions. For direction, this gave values of ±5° (YL), ±6° (CS and JG), ±7° (DO), ±10° (AK), and ±16° (MP). For hue, this gave ±3° (CS and YL), ±4° (DO), ±5° (JG), ±7° (AK), and ±10° (MP). Observers were required to indicate both hue and direction of the target via a 4AFC response – either Blue/Counter-clockwise (the '7' key), Blue/Clockwise ('9'), Pink/Counter-clockwise ('1'), or Pink/Clockwise ('3'). Targets were either presented in isolation or with flankers. For the crowded conditions, two sets of flankers were selected in each domain to be 'strong' and produce maximal effects on threshold elevation and assimilative bias and 'weak' with minimal effects on threshold and bias. We focussed on assimilative bias conditions here to ensure the most consistency between the two domains. This gave flanker directions in the 'strong' condition of ±10° (DO), ±15° (AK, CS, JG, and YL), and ±30° (MP), and in the 'weak' condition of ±165° for 5 observers and ±175°

(AK). For hue, 'strong' flankers were ±15° (AK and YL) or ±30° for the remainder, while 'weak' flankers were ±150° for five observers and ±165° for JG.

These possibilities for crowding strength gave four crowding conditions: the target alone (uncrowded), small target-flanker differences (leading to strong crowding for both colour and motion), large differences for motion and small differences for colour (leading crowding to be weak for motion and strong for colour), and small differences in motion with large differences in colour (strong for motion and weak for colour). Within the three crowding conditions, there were four possible combinations of target and flanker elements in terms of the agreement of the elements in the 2AFC decision space for each feature domain, regardless of the precise values of direction and hue: target and flanker elements could either match in both hue and direction (*both match*; e.g. a clockwise direction for both target and flankers, with all coloured blue), they could differ in their directions only (motion differs; e.g. all blue elements with a clockwise-moving target and CCW flankers), differ in colour only (colour differs), or they could differ in both domains (both differ). The latter conditions were particularly important in separating the predictions of the independent and combined models. The four crowding strength conditions were tested in separate blocks. Each combination of target and flanker elements (in the 2AFC decision space for each feature) was repeated 10 times per block to give a total of 40 trials in the uncrowded condition and 160 trials in the other three. Each block was repeated 6 times to give a total of 3120 trials per observer, plus practice, which observers completed in 3 sessions of around 1 hour each.

Analyses

In Experiments 1 and 2, responses were collected as a proportion of trials in which observers made a counter-clockwise response (corresponding to movement left of upwards in Experiment 1 and stimuli that were pink/purple in Experiment 2). Psychometric functions were fit to this data as a cumulative Gaussian function with three free parameters: the midpoint of the function (at 50%), the slope, and the lapse rate. Shifts in the midpoint were taken as an indication of changes in appearance (i.e. errors of assimilation vs. repulsion). Thresholds were estimated as the difference in direction or hue required to shift performance from the midpoint at 50% to a 75% of CCW responses. Threshold elevation was then obtained by dividing crowded thresholds by uncrowded thresholds. In Experiment 3, we measured percent correct on each of the 2AFC judgements (for colour and motion), which we classified as belonging to one of four quadrants (as indicated on Figure 3): both correct, motion error, colour error, or both error.

Models

Data in Experiments 1 and 2 were fit with a population-coding model based on that developed by Harrison & Bex[26]. In order to capture the pattern of biases and threshold elevation in Experiment 1, the model for motion crowding had 9 free parameters, further details of which are described in the supplementary materials (see Supplementary Figure 1). Because repulsion was not present in the data for Experiment 2, inhibitory components of the model were set to zero, leaving 5 effective free parameters to account for colour crowding. The final best-fitting parameters for both models are shown in Supplementary Table 1, and final outputs shown alongside the data in Figures 1 and 2.

In Experiment 3, two models were developed to contrast between combined and independent mechanisms of crowding. The independent model consisted of population responses to motion and colour, generated for both target and flanker elements and combined according to separate weighting fields for both motion and colour. The combined model was otherwise identical except for the use of a common weights for the two dimensions. Properties of these populations were carried forward from Experiments 1 and 2 for each of the two populations (i.e. the standard deviation of detector tuning functions, as well as the peak height and standard deviation of the positive weighting field). This left 3 free parameters in both models (as described in the Supplementary Materials), with the best-fitting parameters shown in Supplementary Table 2 and the final output shown in Figure 3 (right column). Alternative versions of these models were also developed, as discussed in the Supplementary Materials (see Supplementary Figures 2 and 3, with best-fitting parameters for each model shown in Supplementary Tables 3 and 4).

Acknowledgements

Funded by a Career Development Award (MR/K024817/1) to JG from the UK Medical Research Council. Thanks to Alexandra Kalpadakis-Smith, Samuel Solomon, and Christoph Witzel for helpful comments. Aspects of this work were presented to the Vision Sciences Society[75].

References

1 Whitney, D. & Levi, D. M. Visual crowding: A fundamental limit on conscious perception and object recognition. *Trends in Cognitive Sciences* **15**, 160-168 (2011).

2 Levi, D. M. Crowding - An essential bottleneck for object recognition: A mini-review. *Vision Research* **48**, 635-654 (2008).

3 Bouma, H. Interaction effects in parafoveal letter recognition. *Nature* **226**, 177-178 (1970).

4 Wilkinson, F., Wilson, H. R. & Ellemberg, D. Lateral interactions in peripherally viewed texture arrays. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision* **14**, 2057-2068 (1997).

5 van den Berg, R., Roerdink, J. B. T. M. & Cornelissen, F. W. On the generality of crowding: Visual crowding in size, saturation, and hue compared to orientation. *Journal of Vision* **7**, 1-11 (2007).

6 Kennedy, G. J. & Whitaker, D. The chromatic selectivity of visual crowding. *Journal of Vision* **10**, 1-13, doi:10.1167/10.6.15 (2010).

7 Bex, P. J. & Dakin, S. C. Spatial interference among moving targets. *Vision Research* **45**, 1385-1398 (2005).

8 Butler, T. W. & Westheimer, G. Interference with stereoscopic acuity: Spatial, temporal, and disparity tuning. *Vision Research* **18**, 387-392 (1978).

9 Astle, A. T., McGovern, D. P. & McGraw, P. V. Characterizing the role of disparity information in alleviating visual crowding. *Journal of Vision* **14**, 1-14, doi:10.1167/14.6.8 (2014).

10 Põder, E. & Wagemans, J. Crowding with conjunctions of simple features. *Journal of Vision* **7**, 1-12 (2007).

11 Greenwood, J. A., Bex, P. J. & Dakin, S. C. Crowding follows the binding of relative position and orientation. *Journal of Vision* **12(3):18**, 1-20 (2012).

12 Kooi, F. L., Toet, A., Tripathy, S. P. & Levi, D. M. The effect of similarity and duration on spatial interaction in peripheral vision. *Spatial Vision* **8**, 255-279 (1994).

13 Greenwood, J. A., Bex, P. J. & Dakin, S. C. Positional averaging explains crowding with letter-like stimuli. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 13130-13135 (2009).

14 Greenwood, J. A., Bex, P. J. & Dakin, S. C. Crowding changes appearance. *Current Biology* **20**, 496-501 (2010).

15 Parkes, L., Lund, J., Angelucci, A., Solomon, J. A. & Morgan, M. Compulsory averaging of crowded orientation signals in human vision. *Nature Neuroscience* **4**, 739-744 (2001).

16 Mareschal, I., Morgan, M. J. & Solomon, J. A. Cortical distance determines whether flankers cause crowding or the tilt illusion. *Journal of Vision* **10(8):13**, 1-14, doi:10.1167/10.8.13 (2010).

17 Livingstone, M. & Hubel, D. H. Segregation of form, color, movement, and depth: anatomy, physiology, and perception. *Science* **240**, 740-749 (1988).

18 Ungerleider, L. G. & Mishkin, M. in *Analysis of visual behavior* (eds D.J. Ingle, M.A. Goodale, & R.J.W. Mansfield) 549-586 (MIT Press, 1982).

19 Intriligator, J. & Cavanagh, P. The Spatial Resolution of Visual Attention. *Cognitive Psychology* **43**, 171-216 (2001).

20 Strasburger, H. Unfocussed spatial attention underlies the crowding effect in indirect form vision. *Journal of Vision* **5**, 1024-1037 (2005).

21 He, S., Cavanagh, P. & Intriligator, J. Attentional resolution and the locus of visual awareness. *Nature* **383**, 334-337 (1996).

22 Herzog, M. H. & Manassi, M. Uncorking the bottleneck of crowding: a fresh look at object recognition. *Current Opinion in Behavioral Sciences* **1**, 86-93 (2015).

23 Kane, D., Bex, P. J. & Dakin, S. C. The aperture problem in contoured stimuli. *Journal of Vision* **9(10):13**, 1-17 (2009).

24 Adelson, E. H. & Movshon, J. A. Phenomenal coherence of moving visual patterns. *Nature* **300**, 523-525 (1982).
25 Ester, E. F., Zilber, E. & Serences, J. T. Substitution and pooling in visual crowding induced by similar and dissimilar distractors. *Journal of vision* **15**, 1-12 (2015).
26 Harrison, W. J. & Bex, P. J. A Unifying Model of Orientation Crowding in Peripheral Vision. *Current Biology* **25**, 3213-3219, doi:10.1016/j.cub.2015.10.052 (2015).
27 Glen, J. C. & Dakin, S. C. Orientation-crowding within contours. *Journal of vision* **13**, 1-11 (2013).
28 Blakemore, C., Carpenter, R. H. S. & Georgeson, M. A. Lateral inhibition between orientation detectors in the human visual system. *Nature* **228**, 37-39 (1970).
29 Allman, J., Miezin, F. M. & McGuinness, E. Direction-and velocity-specific responses from beyond the classical receptive field in the middle temporal visual area (MT). *Perception* **14**, 105-126 (1985).
30 De Valois, R. L., Yund, E. W. & Hepler, N. The orientation and direction selectivity of cells in macaque visual cortex. *Vision Research* **22**, 531-544 (1982).
31 Albright, T. D. Direction and orientation selectivity of neurons in visual area MT of the macaque. *Journal of Neurophysiology* **52**, 1106-1130 (1984).
32 Derrington, A. M., Krauskopf, J. & Lennie, P. Chromatic mechanisms in lateral geniculate nucleus of macaque. *The Journal of Physiology* **357**, 241-265 (1984).
33 Witzel, C. & Gegenfurtner, K. R. Categorical sensitivity to color differences. *Journal of Vision* **13**, 1-33, doi:10.1167/13.7.1 (2013).
34 Solomon, S. G. & Lennie, P. The machinery of colour vision. *Nature Reviews Neuroscience* **8**, 276-286 (2007).
35 Lennie, P., Krauskopf, J. & Sclar, G. Chromatic mechanisms in striate cortex of macaque. *Journal of Neuroscience* **10**, 649-669 (1990).
36 Emery, K. J., Volbrecht, V. J., Peterzell, D. H. & Webster, M. A. Variations in normal color vision. VI. Factors underlying individual differences in hue scaling and their implications for models of color appearance. *Vision Research* **141**, 51-65 (2017).
37 Klauke, S. & Wachtler, T. "Tilt" in color space: hue changes induced by chromatic surrounds. *Journal of Vision* **15(13):17**, 1-11 (2015).
38 Marshak, W. & Sekuler, R. Mutual repulsion between moving visual targets. *Nature* **205**, 1399-1401 (1979).
39 Kim, J. & Wilson, H. R. Motion integration over space: interaction of the center and surround motion. *Vision Research* **37**, 991-1005 (1997).
40 Redies, C. & Spillmann, L. The neon color effect in the Ehrenstein illusion. *Perception* **10**, 667-681 (1981).
41 Yashar, A., Wu, X., Chen, J. & Carrasco, M. Crowding and binding: Not all feature-dimensions behave equally [Abstract]. *Journal of Vision* **17**, 374-374 (2017).
42 Watt, R. J., Morgan, M. J. & Ward, R. M. The use of different cues in vernier acuity. *Vision Research* **23**, 991-995 (1983).
43 Edwards, M. & Badcock, D. R. Global-motion perception: Interaction of chromatic and luminance signals. *Vision Research* **36**, 2423-2431 (1996).
44 Sincich, L. C. & Horton, J. C. The circuitry of V1 and V2: Integration of color, form, and motion. *Annual Review of Neuroscience* **28**, 303-326 (2005).
45 Cheadle, S. W. & Zeki, S. Masking within and across visual dimensions: Psychophysical evidence for perceptual segregation of color and motion. *Visual Neuroscience* **28**, 445-451 (2011).
46 Fougnie, D. & Alvarez, G. A. Object features fail independently in visual working memory: Evidence for a probabilistic feature-store model. *Journal of Vision* **11(12):3**, 1-12, doi:10.1167/11.12.3 (2011).
47 Treisman, A. Features and objects: The fourteenth Bartlett memorial lecture. *The Quarterly Journal Of Experimental Psychology* **40**, 201-237 (1988).
48 Wolfe, J. M. & Horowitz, T. S. What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience* **5**, 495-501 (2004).
49 McLeod, P., Driver, J. & Crisp, J. Visual search for a conjunction of movement and form is parallel. *Nature* **332**, 154-155 (1988).

50 Balas, B., Nakano, L. & Rosenholtz, R. A summary-statistic representation in peripheral vision explains visual crowding. *Journal of Vision* **9(12): 13**, 1-18 (2009).
51 Freeman, J. & Simoncelli, E. P. Metamers of the ventral stream. *Nature Neuroscience* **14**, 1195-1201 (2011).
52 Rosenholtz, R. Capabilities and limitations of peripheral vision. *Annual Review of Vision Science* **2**, 437-457 (2016).
53 Saarela, T. P., Sayim, B., Westheimer, G. & Herzog, M. H. Global stimulus configuration modulates crowding. *Journal of Vision* **9**, 1-11 (2009).
54 Rosen, S. & Pelli, D. G. Crowding by a repeating pattern. *Journal of Vision* **15(6):10**, 1-9 (2015).
55 Sayim, B., Westheimer, G. & Herzog, M. H. Gestalt Factors Modulate Basic Spatial Vision. *Psychological Science* **21**, 641-644, doi:10.1177/0956797610368811 (2010).
56 Manassi, M., Sayim, B. & Herzog, M. H. When crowding of crowding leads to uncrowding. *Journal of Vision* **13**, 10-10 (2013).
57 Millin, R., Arman, A. C., Chung, S. T. L. & Tjan, B. S. Visual crowding in V1. *Cerebral Cortex* **24**, 3107-3115 (2013).
58 Kwon, M., Bao, P., Millin, R. & Tjan, B. S. Radial-tangential anisotropy of crowding in the early visual areas. *Journal of Neurophysiology* **112**, 2413-2422 (2014).
59 Bi, T., Cai, P., Zhou, T. & Fang, F. The effect of crowding on orientation-selective adaptation in human early visual cortex. *Journal of Vision* **9**, 13-13 (2009).
60 Motter, B. C. Central V4 Receptive Fields Are Scaled by the V1 Cortical Magnification and Correspond to a Constant-Sized Sampling of the V1 Surface. *Journal of Neuroscience* **29**, 5749-5757, doi:10.1523/jneurosci.4496-08.2009 (2009).
61 Anderson, E. J., Dakin, S. C., Schwarzkopf, D. S., Rees, G. & Greenwood, J. A. The neural correlates of crowding-induced changes in appearance. *Current Biology* **22**, 1199-1206 (2012).
62 Seymour, K., Clifford, C. W. G., Logothetis, N. K. & Bartels, A. The coding of color, motion, and their conjunction in the human visual cortex. *Current Biology* **19**, 177-183 (2009).
63 Manassi, M. & Whitney, D. Multi-level Crowding and the Paradox of Object Recognition in Clutter. *Current Biology* **28**, R127-R133 (2018).
64 Louie, E. G., Bressler, D. W. & Whitney, D. Holistic crowding: Selective interference between configural representations of faces in crowded scenes. *Journal of Vision* **7(2):24**, 1-11 (2007).
65 Kalpadakis-Smith, A. V., Goffaux, V. & Greenwood, J. A. Crowding for faces is determined by visual (not holistic) similarity: Evidence from judgements of eye position. *Scientific Reports* **8**, 1-14, doi:10.1038/s41598-018-30900-0 (2018).
66 Clifford, C. W. G., Wenderoth, P. & Spehar, B. A functional angle on some after-effects in cortical vision. *Proceedings of the Royal Society of London. B. Biological Sciences* **267**, 1705-1710 (2000).
67 Webster, M. A. Visual adaptation. *Annual review of vision science* **1**, 547-567 (2015).
68 Clifford, C. W. G. & Rhodes, G. *Fitting the mind to the world: Adaptation and after-effects in high-level vision*. (Oxford University Press, 2005).
69 Gheri, C., Morgan, M. J. & Solomon, J. A. The relationship between search efficiency and crowding. *Perception* **36**, 1779-1787 (2007).
70 Scolari, M., Kohnen, A., Barton, B. & Awh, E. Spatial attention, preview, and popout: Which factors influence critical spacing in crowded displays? *Journal of Vision* **7**, 1-23, doi:10.1167/7.2.7 (2007).
71 Ishihara, S. *Tests for color-blindness*. (Hongo Harukicho, 1914).
72 Brainard, D. H. The psychophysics toolbox. *Spatial Vision* **10**, 433-436 (1997).
73 Pelli, D. G. The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision* **10**, 437-442 (1997).
74 Pelli, D. G., Palomares, M. & Majaj, N. J. Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *Journal of Vision* **4**, 1136-1169 (2004).
75 Greenwood, J. A. & Parsons, M. J. Dissociable effects of crowding for judgements of colour and motion [Abstract]. *Journal of Vision* **16**, 234-234, doi:10.1167/16.12.234 (2016).

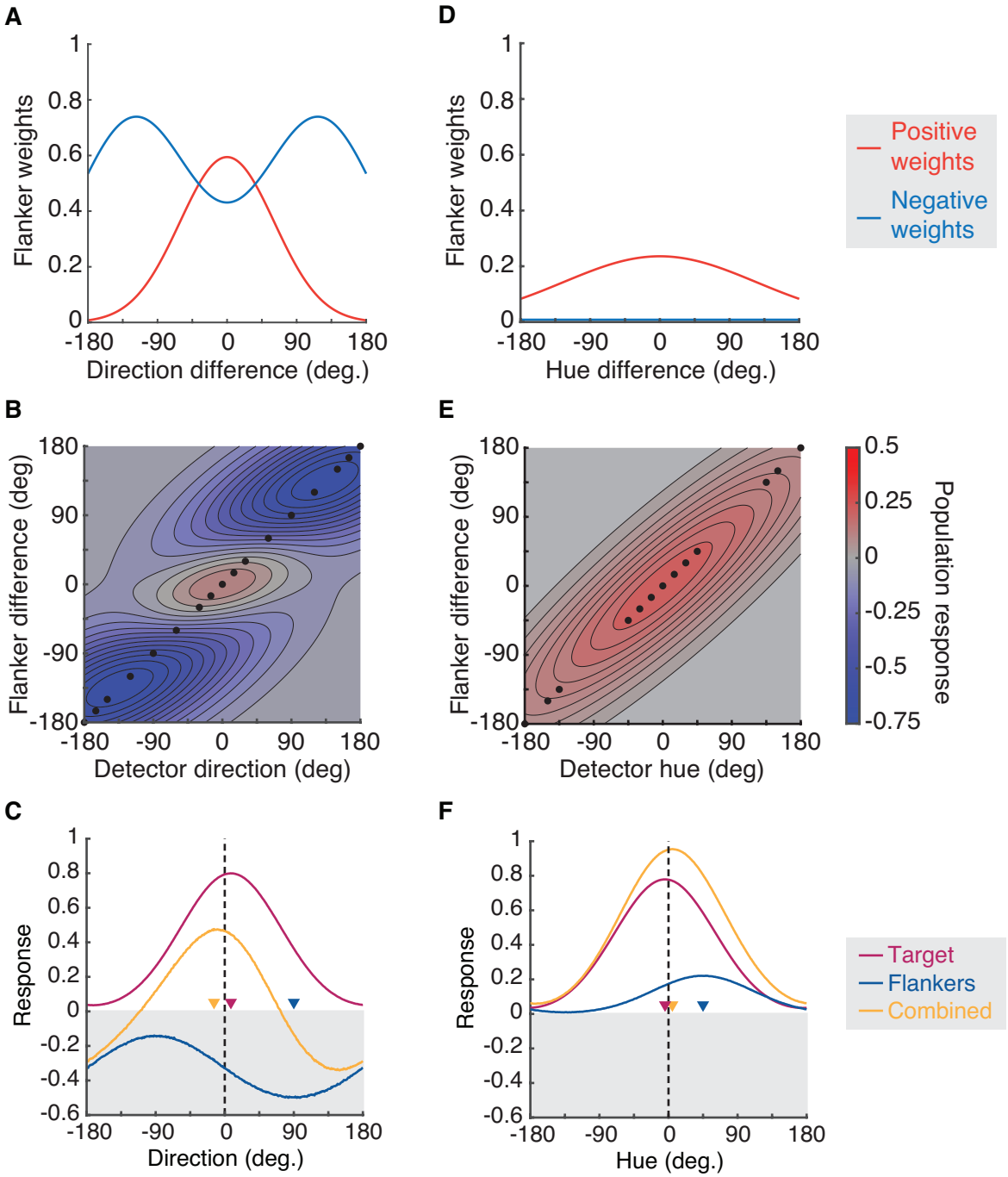
Supplementary Materials for:
Dissociable effects of visual crowding on the perception of colour and motion

John A. Greenwood & Michael J. Parsons
Experimental Psychology, University College London

Population models for the crowding of colour and motion
As outlined in the main manuscript, data from Experiments 1 and 2 were fit with two population-coding models similar to that developed for the crowding of orientation by Harrison & Bex[1]. This approach characterises crowding as the weighted combination of population responses to the target and flanker elements, which has previously been found to reproduce the systematic errors that arise from crowding, including both averaging[2,3] and substitution[4] errors.

Here we sought to extend this model to the domains of motion and colour perception. To replicate the results of Experiment 1, we simulated a population of 361 direction-selective neurons, each with a Gaussian profile of responses to direction, as with those found in cortical areas V1[5] and MT/V5[6]. Each of these detectors had a distinct preferred direction at one of the integer directions from ±180° around upwards, with a peak response of 1 and a standard deviation as the first free parameter. Given the presence of repulsion in our data, we followed recent crowding models[7], models of the tilt illusion[8], and the physiology of MT/V5 neurons[9], by adding inhibitory surrounds to the population response. This was implemented as a Gaussian distribution, with a peak of 1 and a standard deviation as the second free parameter, which was subtracted from the excitatory Gaussian response. For simplicity, this inhibition was applied only to the population response to the flankers, and not to the target response. The result was a population response to the target direction, and a separate population response to the flanker direction. Responses in both cases were corrupted by Gaussian noise, with a magnitude set by one free parameter for the target (detector noise) and another for the flanker distribution (late noise).

As in the model of Harrison & Bex[1], population responses to the target and flanker direction were combined according to a weighting field. In their implementation, this field allowed the decrease in crowding with increasing target-flanker distance. Target-flanker distance was fixed in our study, though we utilise this concept to allow the decrease in crowding with increasing target-flanker dissimilarity[10-13] in direction (following suggestions that both of these properties may in fact manipulate the cortical distance between target and flanker representations[13]). Because our population used both positive and negative components, two weighting fields were applied separately. The first was a Gaussian distribution centred on a target-flanker direction difference of 0, with peak and bandwidth values each set by free parameters. The second was a bimodal Gaussian distribution with peak, peak separation, and bandwidth values as additional free parameters. The shape of this function allowed for the peak in inhibitory interactions at large target-flanker differences where repulsion effects were predominant over assimilation. The best-fitting weighting fields are plotted in Supplementary Figure 1A, which plot the change in flanker weights as a function of the target-flanker difference in direction. The corresponding target weight was always 1 minus the flanker weight.



Supplementary Figure 1. Model characteristics for the population-pooling models of Experiments 1 (left column) and 2 (right). **A.** Weighting fields for direction, plotted separately for the positive (red) and negative (blue) population responses. Flanker weights are plotted as a function of the target-flanker difference in direction. **B.** The combination of flanker weights and the population response to the flanker direction, plotted as a function of the preferred direction of each detector on the x-axis and the target-flanker difference in direction on the y-axis. Flanker values tested in Experiment 1 are shown as black points. Red values indicate a predominance of positive population responses, blue areas indicate a predominance of inhibition, and grey indicates no response. **C.** Example population responses to the target (red), flankers (blue), and combined response (yellow) for a target moving 8° counter-clockwise from upwards and flankers moving 90° counter-clockwise, plotted as a function of the preferred direction of each detector on the x-axis. The veridical values of the target and flanker directions are shown as red and blue triangles, with the peak response of the combined distribution shown as a yellow triangle. This combination gives an error of repulsion. **D.** Weighting fields for hue in Experiment 2, plotted with conventions as in panel A. **E.** The combination of flanker weights and the population response to hue, plotted as in panel B. **F.** Example population responses for a target with a hue -4.5° clockwise from the blue/purple decision boundary (blue in appearance) and flankers 45° counter-clockwise (pink in appearance), which gives an error of assimilation. Plotting conventions are as in panel C.

For a given trial of the simulated experiment, flanker weight values were first selected according to the difference between target and flanker directions. These values were then used as multipliers for the positive and negative components of the population response to the flanker direction. This combination of responses and weighting values for the population response to flankers is plotted in Supplementary Figure 1B for a continuous range of target-flanker differences (flanker values tested in Experiment 1 are shown with black points). Responses are plotted as a function of the preferred direction of each detector in the population on the x-axis against the target-flanker difference on the y-axis. The population response to a single flanker direction can be seen by taking a horizontal slice across the plot, with red areas indicating a predominance of positive flanker responses and blue areas indicating a predominance of inhibition. Note that small target-flanker differences near to the 0° decision boundary tend to produce predominantly positive flanker responses, whereas larger target-flanker differences tend towards inhibition. For each of the positive and negative weighting fields (each ranging from 0-1), the remainder of the weight (1 minus the flanker weight) was applied to the target population response. In the case of the negative component, the target population response was set to 0, as above. With these weights applied, the combined population response was generated by summing the responses to target and flanker directions. The perceived direction of the model on each trial was then taken as the peak response in the combined population distribution.

Example population distributions for a target moving 8° counter-clockwise from upwards and flankers moving 90° counter-clockwise are shown in Supplementary Figure 1C. These distributions (averaged across 1024 trials) have the weights applied to the target and flanker distributions (red and blue lines, respectively). The combined sum is shown in yellow. The veridical values of the target and flankers are shown as triangles. Here it can be seen that the target and flanker directions are both counter-clockwise, yet the peak response for the combined distribution lies on the clockwise side at -9° (yellow triangle) to produce an error of repulsion due to the predominance of inhibition from the flankers on the counter-clockwise side of the population. In contrast, target-flanker combinations where the population response to the flankers was predominantly positive tended to induce assimilation effects by shifting the peak of the combined response to intermediate values between the target and flanker directions (see the demonstration for colour below).

The perceived target direction was derived from the peak response of this combined population response distribution on each of the simulated trials, with the sign of this response used to determine the 2AFC (CW/CCW of vertical) response. Target and flanker direction conditions were identical to those of Experiment 1, with 1024 trials per condition in the simulation. As with the behavioural responses, the percent CCW was then computed for each target direction in each flanker direction condition, with psychometric functions fit to determine midpoint and threshold values. The squared difference between these midpoint and threshold values was then taken from the mean behavioural data of Experiment 1 to give an error term. The best-fitting parameters (for the above 9 free parameters) were determined first using a coarse grid search through the parameter space. A coarse grid search was first run across all possible combinations of parameters to find the least squared error, which was then used as the starting point for a fine fitting procedure using the fminsearch function in MATLAB. The best-fitting parameters derived from this procedure are shown in Supplementary Table 1, and the output of the model is plotted against the data in Figure 1 of the main manuscript.

Supplementary Table 1. Best-fitting free parameter values for the population-based crowding model for motion (Experiment 1) and colour (Experiment 2).

Parameter	Motion model (Experiment 1)	Colour model (Experiment 2)
Detector $SD_{positive}$	65.3691	65.2458
Detector $SD_{negative}$	85.4421	0
Detector noise	0.0153	0.0044
Weight field $Peak_{positive}$	0.5943	0.2358
Weight field $Peak_{negative}$	0.7337	0
Weight field $SD_{positive}$	61.0577	124.6079
Weight field $SD_{negative}$	76.2126	0
Weight field $\Delta\theta_{negative}$	238.5886	0
Late noise	0.1012	0.0802

A similar model was developed to account for the pattern of responses to the colour task of Experiment 2. Because repulsion was not present in the data obtained from this study, all inhibitory components of this model (in the flanker response and weighting fields in particular) were set to zero, leaving 5 free parameters. The structure of the model was otherwise identical to that for motion, with a population of 361 hue-selective neurons with a Gaussian profile of responses to the hue angle in DKL colour space[14-16]. Each detector had a preferred hue angle at one of the integer values in the space with a Gaussian tuning function on either side. This is consistent with suggestions from both physiological[17] and psychophysical results[18], and is used here for ease of comparison across dimensions.

The best-fitting weighting field for colour is shown in Supplementary Figure 1D. Notice that the peak of the weighting field is considerably lower than that for direction, with a broad shallow pattern of tuning. The reduction in crowding for colour is consistent with prior work on the strength of crowding across visual dimensions[19]. The absence of inhibitory weighting fields meant that the sole effect of crowding in the domain of colour was to induce assimilative errors. This can be seen with the combination of the weighting field and flanker population responses across a range of flanker hue angles plotted in Supplementary Figure 1E. As before these values were produced by combining the flanker response to a continuous range of target-flanker differences with the weighting field values for those same target-flanker differences. Responses are plotted as a function of the preferred hue of each detector in the population on the x-axis against the target-flanker difference on the y-axis, with the flanker values tested in Experiment 2 presented as black points. Red areas of response indicate a predominance of positive flanker responses, which occurred for a wider range of flankers in this experiment.

As with motion, the combined population response was generated by summing the responses to target and flanker directions. Example population distributions for a target with a hue angle -4.5° clockwise from the blue/purple decision boundary, and flankers with a 30° counter-clockwise hue are shown in Supplementary Figure 1F. When combined (yellow line) the assimilative nature of crowding can be seen – here the peak response for the combined distribution lies on the counter-clockwise side at 2° (yellow triangle). The best-fitting parameters (for the 5 free parameters) were determined using the same procedure as for motion, and are presented in Supplementary Table 1. The output of this best-fitting model is plotted against the data in Figure 2 of the main manuscript.

Independent and combined population models

In Experiment 3, two models were developed to compare the predictions of combined and independent mechanisms of crowding. The independent model consisted of population responses to motion and colour, generated for both target and flanker elements and combined according to

separate weighting fields for both motion and colour. These separate weighting fields allowed for crowding to occur for one dimension (with small target-flanker differences, e.g. in colour) and not in the other (with larger target-flanker differences for a given dimension, e.g. in motion). The combined model was otherwise identical except for the use of common weights for the two dimensions. In this case, if crowding occurred for one dimension it necessarily occurred for both. For ease of comparison to the combined model, only the positive weighting field and excitatory population responses were used in both models.

The majority of properties of the populations in these models were carried forward from Experiments 1 and 2. This included the standard deviation of detector tuning functions, as well as the peak height and standard deviation of the positive weighting field, as in Supplementary Table 1. This left 3 free parameters in both models: early noise for direction, early noise for hue, and the combined late noise parameter for both dimensions. As in the first two experiments, we simulated the same combination of conditions as performed by observers. Because the precise direction and hue values varied between participants in this experiment (see Methods in the main manuscript), we used the modal value for each. This gave a target direction difference of ±8° and a target hue angle difference of ±5° around the decision boundary in each case. For direction, flankers differed by ±15° in the 'strong motion crowding' conditions and ±165 in the 'weak motion crowding' condition. Hue angles in the flankers were ±30° and ±150°. Each trial simulated the population response to target and flanker values for both motion and colour, as in the models used for the first two experiments.

For the independent model, two separate weighting fields (identical to those of the first two experiments) were used to convert the target-flanker differences in direction and hue into flanker weights. These were then applied to generate a combined population response. The peak response was then used to determine whether responses would be CW/CCW for motion and likewise for hue angle, and percent correct for each judgement was determined. For the combined model, the same procedure was followed (including the two weighting fields), though the same weight value was applied to the flanker response in both colour and motion (with the identical inverse value applied to target responses in each case). Weight values were taken as the maximum value obtained from the weight fields for colour and motion – in this way, if crowding was strong in either dimension it was strong in both (see below for an alternative instantiation of the model where the minimum weights were instead taken).

The 3 free parameters for the independent and combined models were fit by determining the least-squared error between the percent correct scores for motion and colour in each of the four flanker conditions (uncrowded, strong motion + strong colour, weak motion + strong colour, and strong motion + weak colour). Performance was simulated with 1024 trials per point. As before, a coarse grid search was conducted through the parameter space prior to a fine fitting procedure. Final parameters shown in Supplementary Table 2. The output of these best-fitting models is plotted in Figure 3 of the main manuscript.

Supplementary Table 2. Best-fitting free parameter values for the independent and combined models of crowding used to simulate the data of Experiment 3.

Parameter	Independent model	Combined model
Direction noise	0.0256	0.0262
Colour noise	0.0059	0.0041
Late noise	0.0360	0.0908

Alternative versions of the independent and combined population models

In addition to the combined and independent population models of crowding used above to simulate the results of Experiment 3, we also examined the output of a range of alternative versions of these models that could be argued to more closely capture the observed pattern of data. Of particular

interest is whether the combined model could be modified in some way to outperform the independent model. Here we show this to be unlikely.

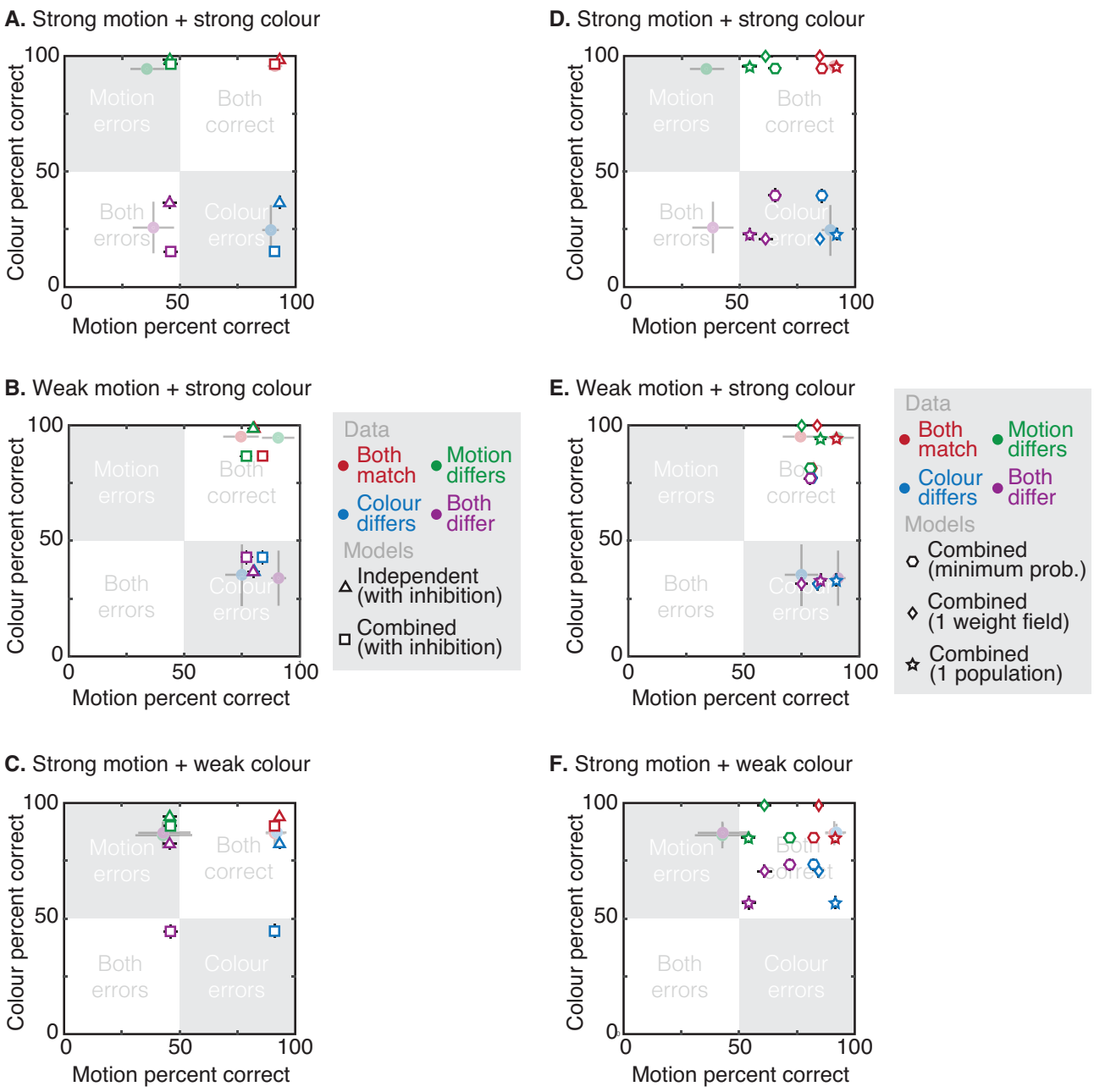
Independent and combined mechanisms with inhibitory interactions

The independent and combined models used to simulate the conjoint colour and motion judgements of Experiment 3 (presented in Figure 3 of the manuscript, with the parameters in Supplementary Table 2 above) did not involve any inhibitory mechanisms in either the population characteristics or the weighting field. Because Experiment 2 revealed purely assimilative effects for the crowding of colour, these negative parameters were set to zero in the population model used to simulate the effect of crowding on colour. To reduce the differences between the independent and combined models in Experiment 3, we therefore removed these parameters from models discussed in the main manuscript, as above. It is however possible that the independent model in particular may perform better with the inclusion of these negative population interactions, given the patterns of repulsion observed for motion in Experiment 1. We therefore reintroduced these values to the population for motion (with the parameters shown in Supplementary Table 1) and determined the best-fitting free parameters for direction noise, colour noise and late noise in both the independent and combined models, using the same fitting procedure as described above. The resulting parameters are shown in Supplementary Table 3.

Supplementary Table 3. Best-fitting free parameter values for the alternative versions of the independent and combined models of crowding used to simulate the data of Experiment 3 with inhibitory parameters for the direction-selective population responses.

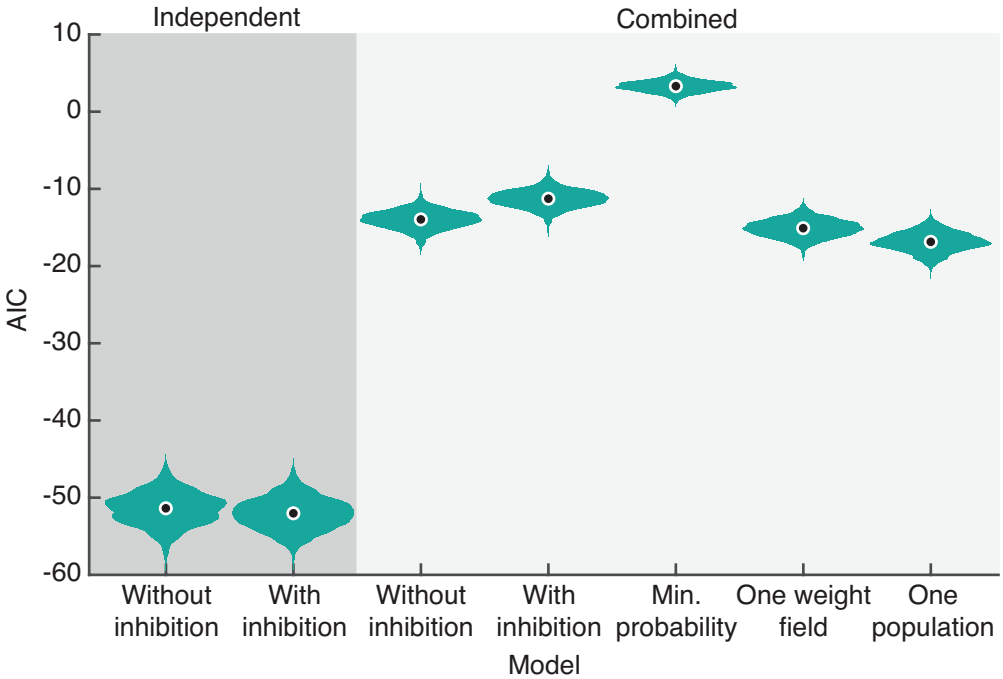
Parameter	Independent model (with inhibition)	Combined model (with inhibition)
Direction noise	0.0310	0.0215
Colour noise	0.0031	0.0116
Late noise	0.0444	0.0623

The final output of these best-fitting models is shown in the left column of Supplementary Figure 2 (panels A-C), plotted with conventions as in Figure 3 of the manuscript. The data is shown with reduced opacity for reference to the two models. Here it can be seen that both models perform similarly to their non-inhibitory counterparts – once again, both models are capable of reproducing errors when crowding was strong for both motion and colour, and similarly when crowding was weak for motion. However, when crowding was strong for motion and weak for colour, the combined model incorrectly predicts errors in both feature domains when the target and flanker differ in both, and colour errors when the target and flanker differ in colour. As a result, the independent model again outperforms the combined model. In order to quantify these differences, for each of these models, the output of 1000 simulations was generated and subtracted from the mean percent correct data from Experiment 3 to obtain squared error values. The mean squared error for the Independent model with inhibition was 0.1074, an improvement from the 0.1101 derived from the Independent model shown in Figure 3. In this case, the fit improves slightly for the model with inhibitory parameters because it allows for a small repulsive effect in judgements for motion when the target-flanker direction differences are large, as we saw in the data of Experiment 1. In contrast, the addition of inhibitory parameters to the combined model produced a squared error value of 0.5147, worse than the value of 0.4641 reported for the combined model shown in Figure 3 of the manuscript.



Supplementary Figure 2. Alternative model simulations for the dual crowding experiment. As in Figure 3 of the main manuscript, each panel plots the percent correct for the target direction of motion on the x-axis and the target hue on the y-axis. With this we can categorise responses as being correct on both dimensions, wrong on motion or colour, or wrong on both, as indicated by the quadrant demarcation. We further divide conditions into 4 target-flanker match cases: where the 2AFC sign in each dimension matches for both (red points), where the motion differs (green), the colour alone differs (blue), or both differ (purple). Data points are shown in all cases with reduced opacity, and as the mean ±1 SEM. **A.** Data for the condition when flankers were selected to produce strong crowding on both dimensions (i.e. with small target-flanker differences in each), along with the output of the alternative model simulations with inhibitory parameters using the best-fitting parameters for both. **B.** Data and simulations for the condition with a large target-flanker difference for motion (which induces weak crowding) and a small target-flanker difference for colour (which induces strong crowding). **C.** Data and simulations for the condition with a large target-flanker difference for colour (weak crowding) and a small target-flanker difference for direction (strong crowding). **D.** Data and simulations for the strong motion + strong colour condition, here with model outputs plotted for the alternative models with the combined mechanism that either use the minimum probability, a common weighting field, or common population parameters for colour and motion. **E.** Data and model simulations for the alternative combined models, plotted as in panel B. **F.** Plotted as in panel C.

For model comparison with all models described herein, the Akaïke Information Criterion (AIC)[20] was computed and is shown in Supplementary Figure 3, where more negative values indicate better fits to the data. Here again it can be seen that the inclusion of inhibition improved the fit of the independent model slightly relative to the model without inhibition presented in the main manuscript, whereas these same inhibitory parameters impaired the fit of the combined model. This is perhaps unsurprising, given the clear differences in the pattern of biases induced in Experiment 1 for motion and Experiment 2 for colour – clearly distinct population dynamics are at play for these two features, offering further evidence for a dissociation in the crowding of colour and motion.



Supplementary Figure 3. Akaike Information Criterion (AIC) values derived from the best-fitting independent and combined models of crowding. Each distribution shows the AIC value for 1000 simulations of the experiment with each model, where the width of the distribution indicates the frequency of the AIC value and the black point shows the mean AIC value. Two independent models are shown – one is the form shown in Figure 3 of the main manuscript, the second includes inhibitory parameters in the direction-selective population. All 5 combined models show less negative AIC values, indicating worse fits. These include the combined model shown in Figure 3, one with inhibitory parameters, one where the Weighting Field is fit specifically to the data (with 5 free parameters) and one where both the weighting field and bandwidth of the detectors are fit specifically.

Alternative versions of the combined-mechanism model

Although the independent model outperforms the combined model both in the main manuscript and with the inhibitory parameters used above, it is possible that alternative versions of the combined model may improve the performance of this mechanism. This is particularly so given the assumptions made in developing these models. For instance, the model presented in Figure 3 took the maximum weight for the flankers in either motion or colour and applied it to both features. In this case, when crowding occurred strongly for one feature, it necessarily occurred strongly for both. Although it is possible that these choices in the design of the combined model biased the outcome towards the independent model, here we show that this is unlikely to be the case by considering whether alternative versions of the combined mechanism could produce a pattern of performance closer to that observed in Experiment 3.

We first simulated a version of the combined model which took the minimum weight for crowding generated in each of the two feature domains. Here, when crowding was weak for either feature it was weak for both. The model was otherwise identical to that reported for the combined mechanism in the

main manuscript. Parameters were fit as before, with the best-fitting free parameters for this model shown in Supplementary Table 4 (listed as 'Minimum Weight). The output of this best-fitting model is shown in the right column of Supplementary Figure 2 (panels D-F) where it can be seen that the model vastly underperforms relative to the maximum-weight combined model, and even more so than the independent model. Here because the weighting field for colour is lower overall, the pattern of performance tends to be driven by colour. Consequently, crowding is reduced in the case with strong motion and strong colour crowding, pushing the responses of the model into the both correct quadrant when motion differs and into the colour errors quadrant when both differ (top row). When motion crowding is reduced (middle row), the model is nonetheless able to account for the colour errors that result because this weighting field is always lower than that for motion. However, the model fails dramatically when colour crowding is reduced (bottom row), predicting that observers should be correct on both features in all conditions, when in fact motion errors predominate in the responses of the observers. As a result, the AIC values for this model (Supplementary Figure 3) are considerably higher than those of the combined model with the maximum probability, and even further from the performance of the independent models. It is therefore highly unlikely that crowding involves a combined mechanism where crowding is automatically reduced for all dimensions when it is reduced for one.

Supplementary Table 4. Best-fitting free parameter values for the alternative versions of the combined models of crowding used to simulate the data of Experiment 3.

Parameter	Combined (min. probability)	Combined (one weight field)	Combined (one popn.)
Direction noise	0.0336	0.0430	0.0106
Colour noise	0.0104	0.0020	0.0106
Late noise	0.0629	0.0196	0.0881
Weight Field SD	.	40.1060	30.2782
Weight Field Peak	.	0.3029	0.4810
Detector SD	.	.	61.9690

As another alternative, it is possible that the combined mechanism shown in Figure 3 underperformed because the Weighting Field for each dimension had distinct parameters, as carried forward from Experiments 1 and 2 (even though only one was applied to both domains on each trial). It is possible that a common weighting field with specifically fit parameters could improve the performance of this model. We therefore added two additional free parameters, to fit the peak and SD values of a single weighting field that was applied to the judgements of both colour and motion in Experiment 3. The remaining model details were as in the other versions of this model, as with the fitting procedure.

The best-fitting free parameters for this model are shown in Supplementary Table 4 (listed as 'one weight field'), with the resulting output shown in the right column of Supplementary Figure 2. The best-fitting weighting field here is narrower than that for either feature domain in the main model, with an intermediate peak height. As a result, the model underperforms on the motion errors in particular, underestimating the rate of motion errors even in the both strong condition. Although the model is consequently able to capture the shift to colour errors in the weak motion condition, it fails dramatically on the weak colour condition, predicting that crowding should be weak in both dimensions (a consequence of the narrow tuning for the weighting field and the slightly larger values used in the target-flanker direction differences). The error for this model is accordingly large, with the resulting AIC values shown in Supplementary Figure 3. The mean AIC for this model was -15.09, improving slightly on the value of -13.96 for the combined model in the main manuscript. A specifically tailored weighting field can thus improve the fit of the combined model to some extent,

however neither model approaches the fits obtained with the independent model. Errors are still all-or-none, which fails to predict the release from crowding in one domain but not the other that occurs when motion crowding is reduced in comparison to colour, or vice versa.

Finally, following the logic above, it is possible that the combined model could be further improved with the use of a single population that shares the same characteristics for both colour and motion (i.e. a detector population with the same bandwidth and noise values for both colour and motion), as well as a common weighting field. In a sense, we can already rule this out as unlikely given the vastly distinct response profiles produced for the crowding of motion and colour in Experiments 1 and 2 – for the same sized stimuli in the same locations at the same eccentricity, crowding affected motion with a mixture of assimilation and repulsion, while crowding for colour was characterised only by assimilation. Nonetheless, in order to see whether this approach could improve the fits for the combined model, we developed a version of the model with matched population parameters for colour and motion (to give an additional 2 free parameters). In order to find population characteristics that could reasonably account for the results of all 3 experiments, we fit the 5 free parameters of this model (detector noise and bandwidth, weighting field peak and bandwidth, and late noise) to all 3 experiments simultaneously to determine the best-fitting parameters.

The resulting parameters for this model are shown in Supplementary Table 4 (as 'one population'), and the final output of the model in Experiment 3 is shown in the right column of Supplementary Figure 2. Notice that the best-fitting weighting field is again considerably narrower than that obtained for the combined model of the main manuscript. This leads to a similar pattern of performance to that seen for the model with a single weighting field above. Performance in the condition where crowding was strong for both motion and colour tends to underestimate the degree of motion crowding as a result. This becomes particularly problematic in the strong motion condition, where the narrow weighting field gives a strong release from crowding (to give predominantly 'both correct' responses), despite the motion errors that predominate in the observers responses. This model nonetheless gave a slight improvement to the fit of the combined model to the data, with a squared error value of 0.36 and an AIC of -16.86 on average, as shown in Supplementary Figure 3. This is still far off from the performance of either instantiation of the independent model, however.

Altogether, although some of these variants perform marginally better than the combined model presented in the main manuscript, none approaches the performance of the independent model. Given the behavioural evidence presented in the manuscript, where independent crowding effects disrupt the domains of motion and colour, in addition to the results of these simulations whereby a population-coding model of this process could account for this pattern of results when weighting fields for the two dimensions were independent, we argue that crowding disrupts the perception of colour and motion independently.

References

1 Harrison, W. J. & Bex, P. J. A Unifying Model of Orientation Crowding in Peripheral Vision. *Current Biology* **25**, 3213-3219, (2015).

2 Parkes, L., Lund, J., Angelucci, A., Solomon, J. A. & Morgan, M. Compulsory averaging of crowded orientation signals in human vision. *Nature Neuroscience* **4**, 739-744, (2001).

3 Greenwood, J. A., Bex, P. J. & Dakin, S. C. Positional averaging explains crowding with letter-like stimuli. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 13130-13135, (2009).

4 Ester, E. F., Zilber, E. & Serences, J. T. Substitution and pooling in visual crowding induced by similar and dissimilar distractors. *Journal of vision* **15**, 1-12, (2015).

5 De Valois, R. L., Yund, E. W. & Hepler, N. The orientation and direction selectivity of cells in macaque visual cortex. *Vision Research* **22**, 531-544, (1982).

6 Albright, T. D. Direction and orientation selectivity of neurons in visual area MT of the macaque. *Journal of Neurophysiology* **52**, 1106-1130, (1984).

7 Glen, J. C. & Dakin, S. C. Orientation-crowding within contours. *Journal of vision* **13**, 1-11, (2013).

8 Blakemore, C., Carpenter, R. H. S. & Georgeson, M. A. Lateral inhibition between orientation detectors in the human visual system. *Nature* **228**, 37-39, (1970).
9 Allman, J., Miezin, F. M. & McGuinness, E. Direction-and velocity-specific responses from beyond the classical receptive field in the middle temporal visual area (MT). *Perception* **14**, 105-126, (1985).
10 Wilkinson, F., Wilson, H. R. & Ellemberg, D. Lateral interactions in peripherally viewed texture arrays. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision* **14**, 2057-2068, (1997).
11 Kennedy, G. J. & Whitaker, D. The chromatic selectivity of visual crowding. *Journal of Vision* **10**, 1-13, (2010).
12 Kooi, F. L., Toet, A., Tripathy, S. P. & Levi, D. M. The effect of similarity and duration on spatial interaction in peripheral vision. *Spatial Vision* **8**, 255-279, (1994).
13 Mareschal, I., Morgan, M. J. & Solomon, J. A. Cortical distance determines whether flankers cause crowding or the tilt illusion. *Journal of Vision* **10(8):13**, 1-14, (2010).
14 Derrington, A. M., Krauskopf, J. & Lennie, P. Chromatic mechanisms in lateral geniculate nucleus of macaque. *The Journal of Physiology* **357**, 241-265, (1984).
15 Solomon, S. G. & Lennie, P. The machinery of colour vision. *Nature Reviews Neuroscience* **8**, 276-286, (2007).
16 Witzel, C. & Gegenfurtner, K. R. Categorical sensitivity to color differences. *Journal of Vision* **13**, 1-33, (2013).
17 Lennie, P., Krauskopf, J. & Sclar, G. Chromatic mechanisms in striate cortex of macaque. *Journal of Neuroscience* **10**, 649-669, (1990).
18 Emery, K. J., Volbrecht, V. J., Peterzell, D. H. & Webster, M. A. Variations in normal color vision. VI. Factors underlying individual differences in hue scaling and their implications for models of color appearance. *Vision Research* **141**, 51-65, (2017).
19 van den Berg, R., Roerdink, J. B. T. M. & Cornelissen, F. W. On the generality of crowding: Visual crowding in size, saturation, and hue compared to orientation. *Journal of Vision* **7**, 1-11, (2007).
20 Akaike, H. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* **19**, 716–723, (1974).